Concavity of certain matrix trace and norm functions

Fumio Hiai[1]

[1] Tohoku University (Emeritus),
Hakusan 3-8-16-303, Abiko 270-1154, Japan

Abstract

Lieb's concavity/convexity and Epstein's concavity are extended by improving Epstein's method to matrix trace functions in certain general forms, which are further generalized by the majorization method to concavity/convexity of similar functions under symmetric (anti-) norms in place of the trace.

2010 Mathematics Subject Classification: Primary 15A60, 47A30, 47A60

Key words and phrases: Matrices, trace, symmetric norms, symmetric anti-norns, joint concavity, joint convexity, operator mean

Introduction

In the present paper we consider the matrix functions of the following types:

(i) $F(A,B) = \{\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\}^s$,

(ii) $F(A,B) = \{\Phi(A^p)\,\sigma\,\Psi(B^q)\}^s$,

(iii) $F(A) = \Phi(A^p)^s$.

Here, the variables A and B are positive definite matrices, Φ and Ψ are (strictly) positive linear maps between matrix algebras, and p,q,s are real parameters. Moreover, σ in (ii) is an operator mean in the Kubo-Ando sense [16]. For matrix functions F in the above we are interested in the range of the parameters p,q,s (or p,s) for which the function

$$(A,B) \longmapsto \|F(A,B)\| \quad (\text{or} \ \ A \longmapsto \|F(A)\|)$$

is convex for any symmetric norm $\|\cdot\|$ and also for which the function $(A,B) \mapsto \|F(A,B)\|_!$ (or $A \mapsto \|F(A)\|_!$) is concave for any symmetric anti-norm $\|\cdot\|_!$ (and for every Φ,Ψ). Here, the notion of symmetric anti-norms was recently introduced in [6]

[1]E-mail: hiai.fumio@gmail.com

while that of symmetric norms is familiar in matrix analysis (see [5, 14] for example). Indeed, a symmetric anti-norm is a non-negative functional on the positive part of a matrix algebra that is invariant under unitary conjugation and superadditive (opposite to the subadditivity of symmetric norms). This is the reason why we take a symmetric anti-norm for concavity assertions while a symmetric norm is for convexity assertions. It is worthwhile to note that the trace functional is a symmetric norm and a symmetric anti-norm in common.

For instance, when $s = 1$, $\Phi(A) := X^*AX$ and $\Psi = \mathrm{id}$, the function (i) under the trace is

$$(A, B) \longmapsto \mathrm{Tr}\, X^*A^pXB^q,$$

whose joint concavity/convexity is famous as Lieb's concavity/convexity [17]. It is well known that an equivalent reformulation is matrix concavity/convexity of $(A, B) \mapsto A^p \otimes B^q$ due to Ando [1]. When $0 < p \le 1$, $s = 1/p$ and $\Phi(A) := X^*AX$, the function (iii) under the trace is

$$A \longmapsto \mathrm{Tr}\,(X^*A^pX)^{1/p},$$

whose concavity was proved by Epstein [10] by a powerful method using theory of Pick functions (often called Epstein's method). The method was applied in [13] to prove (joint) concavity of the trace functions of types (i)–(iii) under certain respective conditions on p, q, s. The Minkowski type trace function (or the trace function for the matrix power means) $\mathrm{Tr}\,(A^p + B^p)^{1/p}$ was discussed in [8] (also [2, 4]), which is a special case of the function (ii) under the trace where $s = 1/p$, $\Phi = \Psi = \mathrm{id}$ and σ is the arithmetic mean. More recently in [9], Carlen and Lieb extensively developed concavity/convexity properties of the trace functions of the forms $\mathrm{Tr}\,(X^*A^pX)^s$ and $\mathrm{Tr}\,(A^p+B^p)^s$ among others. Furthermore in [15], Jenčová and Ruskai obtained equality conditions for Lieb's concavity/convexity as well as for some related inequalities. In this way, the functions of the above types (i)–(iii) cover many of important cases appearing in the study of concavity/convexity of various matrix trace functions so far.

The present paper is a continuation of [13]. Our strategy here is two-fold. We first polish Epstein's method to extend some known concavity/convexity results for trace functions as much as possible. After this is done we further extend the results with the trace to those with symmetric (anti-) norms by using the majorization method. In Section 1 we treat the function (i) under the trace and prove its joint concavity/convexity under suitable conditions on p, q, s by using Epstein's method. In Section 2 we prove joint concavity/convexity of the function (ii) under symmetric (anti-) norms under suitable conditions on p, q, s. In Section 3 we treat the function (iii) and extend a main convexity result in [9] to the case with symmetric norms. In Section 4 we examine necessary conditions on the parameters p, q, s (or p, s) for concavity/convexity of the relevant functions to hold, and compare them with sufficient conditions obtained in Sections 1–3. Here, Bekjan's idea in [4] (also used in [9]) is of particular use. Since

it does not seem easy to extend the result of Section 1 to functions with symmetric (anti-) norms, we consider, in Section 5, the function (i) with the operator norm and the smallest singular value, which are particular cases of the Ky Fan (anti-) norms. (Note that concavity/convexity under all symmetric (anti-) norms can be reduced to that under all Ky Fan (anti-) norms.)

1 Trace functions of Lieb type

We begin with fixing some common notations. For each $n \in \mathbb{N}$ the $n \times n$ complex matrix algebra is denoted by $\mathbb{M}_n$. We write $\mathbb{M}_n^+ := \{A \in \mathbb{M}_n : A \geq 0\}$, the $n \times n$ positive semidefinite matrices, and $\mathbb{P}_n := \{A \in \mathbb{M}_n : A > 0\}$, the $n \times n$ positive definite matrices. The usual trace on $\mathbb{M}_n$ is denoted by Tr. A linear map $\Phi : \mathbb{M}_n \to \mathbb{M}_m$ is positive if $A \in \mathbb{M}_n^+$ implies $\Phi(A) \in \mathbb{M}_m^+$, and it is strictly positive if $A \in \mathbb{P}_n$ implies $\Phi(A) \in \mathbb{P}_m$. Clearly, a positive linear map $\Phi : \mathbb{M}_n \to \mathbb{M}_m$ is strictly positive if $\Phi(I_n) \in \mathbb{P}_m$, where I_n is the identity of $\mathbb{M}_n$.

In this section we consider joint concavity and convexity of the trace function

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \mathrm{Tr}\left\{\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\right\}^s, \tag{1.1}$$

where Φ and Ψ are (strictly) positive linear maps between matrix algebras. In particular, when $s = 1$, $\Phi(A) := X^*AX$ and $\Psi = \mathrm{id}$, the identity map, the above function is

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_n \longmapsto \mathrm{Tr}\, X^*A^pXB^q, \tag{1.2}$$

for which Lieb's concavity (also convexity) is well known [17] (also [1]).

Throughout the section we assume that $(p, q) \neq (0, 0)$ and $s \neq 0$; otherwise, the function (1.1) is constant. The next theorem extends [13, Theorem 2.3].

Theorem 1.1. *Let $n, m, l \in \mathbb{N}$. Let $\Phi : \mathbb{M}_n \to \mathbb{M}_l$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_l$ be strictly positive linear maps.*

(1) *If either $0 \leq p, q \leq 1$ and $1/2 \leq s \leq 1/(p+q)$, or $-1 \leq p, q \leq 0$ and $1/(p+q) \leq s \leq -1/2$, then the function* (1.1) *is jointly concave.*

(2) *If either $-1 \leq p, q \leq 0$ and $1/2 \leq s \leq -1/(p+q)$, or $0 \leq p, q \leq 1$ and $-1/(p+q) \leq s \leq -1/2$, then the function* (1.1) *is jointly convex.*

Proof. (1) First, assume that $0 \leq p, q \leq 1$ and $1/2 \leq s \leq 1/(p+q)$. Although the proof below is a slight improvement of Epstein's method in the proof of [13, Theorem 2.3], we shall present it in detail while the proofs for other cases in (1) and (2) will only be sketched. Let us first show that the assertion in the case $1/2 \leq s < 1/(p+q)$ follows from that in the case $s = 1/(p+q)$. Indeed, when $1/2 \leq s < 1/(p+q)$ (and also $0 < p, q \leq 1$), one can choose $p' \in [p, 1]$ and $q' \in [q, 1]$ such that $s = 1/(p' + q')$. Let

$A_1, A_2 \in \mathbb{M}_n^+$, $B_1, B_2 \in \mathbb{M}_m^+$, and $0 < \lambda < 1$. Then, since x^α $(x \geq 0)$ with $0 < \alpha \leq 1$ is operator concave as well as operator monotone, we have

$$(\lambda A_1 + (1-\lambda)A_2)^p \geq (\lambda A_1^{p/p'} + (1-\lambda)A_2^{p/p'})^{p'}$$

so that

$$\Phi((\lambda A_1 + (1-\lambda)A_2)^p) \geq \Phi((\lambda A_1^{p/p'} + (1-\lambda)A_2^{p/p'})^{p'}),$$

and similarly

$$\Psi((\lambda B_1 + (1-\lambda)B_2)^q) \geq \Psi((\lambda B_1^{q/q'} + (1-\lambda)B_2^{q/q'})^{q'}).$$

From the assertion in the case $s = 1/(p'+q')$ we have

$$\begin{aligned}
&\mathrm{Tr}\left\{\Phi((\lambda A_1 + (1-\lambda)A_2)^p)^{1/2}\Psi((\lambda B_1 + (1-\lambda)B_2)^q)\Phi((\lambda A_1 + (1-\lambda)A_2)^p)^{1/2}\right\}^s \\
&\quad \geq \mathrm{Tr}\left\{\Phi((\lambda A_1^{p/p'} + (1-\lambda)A_2^{p/p'})^{p'})^{1/2}\Psi((\lambda B_1^{q/q'} + (1-\lambda)B_2^{q/q'})^{q'})\right. \\
&\qquad\qquad \left.\times\, \Phi((\lambda A_1^{p/p'} + (1-\lambda)A_2^{p/p'})^{p'})^{1/2}\right\}^s \\
&\quad \geq \lambda \mathrm{Tr}\left\{\Phi((A_1^{p/p'})^{p'})^{1/2}\Psi((B_1^{q/q'})^{q'})\Phi((A_1^{p/p'})^{p'})^{1/2}\right\}^s \\
&\qquad\quad + (1-\lambda)\mathrm{Tr}\left\{\Phi((A_2^{p/p'})^{p'})^{1/2}\Psi((B_2^{q/q'})^{q'})\Phi((A_2^{p/p'})^{p'})^{1/2}\right\}^s \\
&\quad = \lambda \mathrm{Tr}\left\{\Phi(A_1)^{1/2}\Psi(B_1)\Phi(A_1)^{1/2}\right\}^s + (1-\lambda)\mathrm{Tr}\left\{\Phi(A_2^p)^{1/2}\Psi(B_2^q)\Phi(A_2^p)^{1/2}\right\}^s.
\end{aligned}$$

Therefore, in the following proof we may assume that $s = 1/(p+q)$.

Set $\gamma := p+q \in (0,2]$ and so $s = 1/\gamma$. As in [13] we will use the following notations:

$$\mathbb{C}^+ := \{z \in \mathbb{C} : \mathrm{Im}\, z > 0\},$$

$$\mathcal{I}_n^+ := \{X \in \mathbb{M}_n : \mathrm{Im}\, X > 0\}, \qquad \mathcal{I}_n^- := \{X \in \mathbb{M}_n : \mathrm{Im}\, X < 0\},$$

and

$$\Gamma_{\gamma\pi} := \{re^{i\theta} : r > 0,\ 0 < \theta < \gamma\pi\}.$$

Note that for each $\alpha > 0$ the function x^α $(x > 0)$ has the analytic continuation z^α in $\mathbb{C} \setminus [0,\infty)$ (in particular, in $\mathbb{C}^+$) defined by

$$z^\alpha := r^\alpha e^{i\alpha\theta} \quad \text{for} \quad z = re^{i\theta} \ \ (r > 0,\ 0 < \theta < 2\pi).$$

To obtain the joint concavity result, it suffices to prove that if $A, H \in \mathbb{M}_n$ and $B, K \in \mathbb{M}_m$ are such that $A, B > 0$ and H, K are Hermitian, then

$$\frac{d^2}{dx^2}\mathrm{Tr}\left\{\Phi((A+xH)^p)^{1/2}\Psi((B+xK)^q)\Phi((A+xH)^p)^{1/2}\right\}^s \leq 0$$

for every sufficiently small $x > 0$.

For $z \in \mathbb{C}$ set $X(z) := zA + H$ and $Y(z) := zB + K$. For any $z \in \mathbb{C}^+$, since $X(z) \in \mathcal{I}_n^+$, $Y(z) \in \mathcal{I}_m^+$ and $p, q \in (0,1]$, we can define $X(z)^p$ and $Y(z)^q$ by analytic functional calculus by [13, Lemma 1.1]. Since [13, Lemma 1.2] implies that

$$\operatorname{Im}\Phi(X(z)^p) = \Phi(\operatorname{Im} X(z)^p) > 0, \qquad \operatorname{Im}\Psi(Y(z)^q) = \Psi(\operatorname{Im} Y(z)^q) > 0, \tag{1.3}$$

we have $\Phi(X(z)^p), \Psi(Y(z)^q) \in \mathcal{I}_l^+$ and hence $\Phi(X(z)^p)^{1/2} \in \mathcal{I}_l^+$ is also well defined. Now define

$$F(z) := \Phi(X(z)^p)^{1/2}\Psi(Y(z)^q)\Phi(X(z)^p)^{1/2}, \qquad z \in \mathbb{C}^+, \tag{1.4}$$

which is analytic in $\mathbb{C}^+$. We will prove that

$$\sigma(F(z)) \subset \Gamma_{\gamma\pi} \quad \text{if} \quad z \in \mathbb{C}^+, \tag{1.5}$$

where $\sigma(F(z))$ is the set of the eigenvalues of $F(z)$. To prove this, it suffices to show the following properties:

(a) When $z = re^{i\theta}$ with a fixed $0 < \theta < \pi$, $\sigma(F(z)) \subset \Gamma_{\gamma\pi}$ for sufficiently large $r > 0$.

(b) $\sigma(F(z)) \cap [0, \infty) = \emptyset$ for all $z \in \mathbb{C}^+$.

(c) $\sigma(F(z)) \cap \{re^{i\gamma\pi} : r \geq 0\} = \emptyset$ for all $z \in \mathbb{C}^+$.

In fact, if (1.5) fails to hold for some $z_0 = r_0 e^{i\theta_0} \in \mathbb{C}^+$, then according to (a) and the continuity of the eigenvalues of $F(z)$ we must have $\sigma(F(z)) \cup \partial\Gamma_{\gamma\pi} \neq \emptyset$ for some $z \in \{re^{\theta_0} : r > r_0\}$, which says that (b) or (c) must be violated.

Proof of (a). Define

$$\tilde{F}(z) := z^\gamma \Phi((A + z^{-1}H)^p)^{1/2}\Psi((B + z^{-1}K)^q)\Phi((A + z^{-1}H)^p)^{1/2}, \quad z \in \mathbb{C}^+,$$

which is analytic in $\mathbb{C}^+$ (see [13, Section 1]). Since

$$\begin{aligned}\tilde{F}(x) &= x^\gamma \Phi((A + x^{-1}H)^p)^{1/2}\Psi((B + x^{-1}K)^q)\Phi((A + x^{-1}H)^p)^{1/2} \\ &= \Phi((xA + H)^p)^{1/2}\Psi((xB + K)^q)\Phi((xA + H)^p)^{1/2} = F(x)\end{aligned}$$

for all sufficiently large $x > 0$, we obtain

$$F(z) := z^\gamma \Phi((A + z^{-1}H)^p)^{1/2}\Psi((B + z^{-1}K)^q)\Phi((A + z^{-1}H)^p)^{1/2}, \quad z \in \mathbb{C}^+. \tag{1.6}$$

When $z = re^{i\theta_0}$ with $0 < \theta_0 < \pi$ fixed and $r \to \infty$, note that

$$\sigma\big(\Phi((A + z^{-1}H)^p)^{1/2}\Psi((B + z^{-1}K)^q)\Phi((A + z^{-1}H)^p)^{1/2}\big)$$

converges to $S := \sigma\big(\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\big) \subset (0, \infty)$. Since $(r^\gamma e^{i\gamma\theta_0})^{-1}\sigma(F(re^{i\theta_0}))$ converges to S as $r \to \infty$, we see that (a) holds.

Proof of (b). For any $r \in [0,\infty)$ we have

$$F(z) - rI_l = \Phi(X(z)^p)^{1/2}\big(\Psi(Y(z)^q) - r\Phi(X(z)^p)^{-1}\big)\Phi(X(z)^p)^{1/2}.$$

Since $\Phi(X(z)^p), \Psi(Y(z)^q) \in \mathcal{I}_l^+$ as already mentioned, we see by [13, Lemma 1.1] that

$$\Psi(Y(z)^q) - r\Phi(Y(z)^p)^{-1} \in \mathcal{I}_l^+$$

so that $F(z) - rI_l$ is invertible.

Proof of (c). For any $r \in [0,\infty)$ we have

$$F(z) - re^{i\gamma\pi}I_l = e^{iq\pi}\Phi(X(z)^p)^{1/2}\big(\Psi(e^{-iq\pi}Y(z)^q) - r\Phi(e^{-ip\pi}X(z)^p)^{-1}\big)\Phi(X(z)^p)^{1/2}$$

thanks to $p+q=\gamma$. Since $\Psi(e^{-iq\pi}Y(z)^q) - r\Phi(e^{-ip\pi}X(z)^p)^{-1} \in \mathcal{I}_l^-$ by [13, Lemma 1.2], $F(z) - re^{i\gamma\pi}I_l$ is invertible.

We have shown (1.5). Hence we can define $F(z)^s$ for $z \in \mathbb{C}^+$ by applying the analytic functional calculus by z^s on $\Gamma_{\gamma\pi}$ to $F(z)$. Since $\gamma s = 1$ by assumption, note that z^s maps $\Gamma_{\gamma\pi}$ into $\mathbb{C}^+$. Thus, $F(z)^s$ is an analytic function such that $\sigma(F(z)^s) \subset \mathbb{C}^+$ for all $z \in \mathbb{C}^+$ (see [13, Section 1]). In view of (1.6), one can choose an $R > 0$ so that $F(z)^s$ in $\mathbb{C}^+$ is continuously extended to $\mathbb{C}^+ \cup (R,\infty)$ with

$$F(x)^s = x\big\{\Phi((A+x^{-1}H)^p)^{1/2}\Psi((B+x^{-1}K)^q)\Phi((A+x^{-1}H)^p)^{1/2}\big\}^s, \quad x \in (R,\infty).$$

Since $\mathrm{Tr}\,(F(x)^s) \in \mathbb{R}$ for all $x \in (R,\infty)$, by the reflection principle we obtain a Pick function φ on $\mathbb{C} \setminus (-\infty, R]$ such that $\varphi(x) = \mathrm{Tr}\,(F(x)^s)$ for all $x \in (R,\infty)$. For every $x \in (0, R^{-1})$ we have

$$x\varphi(x^{-1}) = \mathrm{Tr}\,\big\{\Phi((A+xH)^p)^{1/2}\Psi((B+xK)^q)\Phi((A+xH)^p)^{1/2}\big\}^s. \tag{1.7}$$

It thus remains to show that

$$\frac{d^2}{dx^2}(x\varphi(x^{-1})) \le 0, \qquad x \in (0, R^{-1}). \tag{1.8}$$

According to Nevanlinna's theorem for Pick functions (see, e.g., [14, Theorem 2.6.2]), φ admits an integral expression

$$\varphi(z) = a + bz + \int_{-\infty}^{\infty} \frac{1+tz}{t-z}\,d\nu(t), \tag{1.9}$$

where $a \in \mathbb{R}$, $b \ge 0$, and ν is a finite measure on $\mathbb{R}$. Since φ is analytically continued across the interval (R,∞), the measure ν is supported in $(-\infty, R]$. Therefore,

$$x\varphi(x^{-1}) = ax + b + \int_{-\infty}^{R} \frac{x(x+t)}{xt-1}\,d\nu(t), \qquad x \in (0, R^{-1}).$$

Compute

$$\frac{d}{dx}\left(\frac{x(x+t)}{xt-1}\right)=\frac{x^2t-2x-t}{(xt-1)^2},\qquad \frac{d^2}{dx^2}\left(\frac{x(x+t)}{xt-1}\right)=\frac{2(t^2+1)}{(xt-1)^3}\le 0$$

for all $x\in(0,R^{-1})$, and hence (1.8) follows.

The proof for the second case where $-1\le p,q\le 0$ and $1/(p+q)\le s\le -1/2$ can be done similarly to the above but a more convenient way is to replace Φ and Ψ with $\hat\Phi(A):=\Phi(A^{-1})^{-1}$ for $A\in\mathbb{P}_n$ and $\hat\Psi(B):=\Psi(B^{-1})^{-1}$ for $B\in\mathbb{P}_m$, respectively. Although $\hat\Phi$ and $\hat\Psi$ are no longer linear, the above proof can work with $\hat\Phi$ and $\hat\Psi$ in place of Φ and Ψ. Indeed, in the above we only used monotonicity and property (1.3) for Φ,Ψ, which are valid for $\hat\Phi,\hat\Psi$ too. Since

$$\bigl\{\hat\Phi(A^p)^{1/2}\hat\Psi(B^q)\hat\Phi(A^p)^{1/2}\bigr\}^s=\bigl\{\Phi(A^{-p})^{1/2}\Psi(B^{-q})\Phi(A^{-p})^{1/2}\bigr\}^{-s}, \tag{1.10}$$

the second case of (1) immediately follows from the first case for $\hat\Phi$ and $\hat\Psi$.

(2) Assume that $-1\le p,q\le 0$ and $1/2\le s\le -1/(p+q)$. As in the proof of (1) we may assume that $s=-1/(p+q)$. Then the proof is similar to the above (1). In the last part we may use $-F(z)^s$ in place of $F(z)^s$. The argument in the last paragraph of the proof of (1) can also work to prove the second case of (2). □

It is obvious by convergence that in Theorem 1.1 strict positivity of Φ,Ψ is relaxed to the usual positivity and $A,B>0$ is to $A,B\ge 0$ as far as all the parameters p, q and s are non-negative. (Here, for $A\in\mathbb{M}_n^+$ both conventions of A^0 being I_n and of A^0 being the support projection of A are available.) This remark will be available throughout the paper.

2 Norm functions involving operator means

Before going into the main topic of this section we recall symmetric anti-norms introduced in [6]. A norm $\|\cdot\|$ on $\mathbb{M}_n$ is said to be *symmetric* or *unitarily invariant* if $\|UXV\|=\|X\|$ for all $X\in\mathbb{M}_n$ and unitaries $U,V\in\mathbb{M}_n$. On the other hand, a *symmetric anti-norm* $\|\cdot\|_!$ on $\mathbb{M}_n^+$ is a non-negative continuous functional such that

(a) $\|\lambda A\|_!=\lambda\|A\|_!$ for all $A\in\mathbb{M}_n^+$ and all reals $\lambda\ge 0$,

(b) $\|A\|_!=\|UAU^*\|_!$ for all $A\in\mathbb{M}_n^+$ and all unitaries U,

(c) $\|A+B\|_!\ge\|A\|_!+\|B\|_!$ for all A, $B\in\mathbb{M}_n^+$.

We note that a symmetric norm $\|\cdot\|$, when restricted on $\mathbb{M}_n^+$, is also characterized by the same (a), (b) and the double inequality $\|A\|\le\|A+B\|\le\|A\|+\|B\|$ in place of the superadditivity axiom in (c). So, the notion of symmetric anti-norms is a natural superadditive counterpart of that of symmetric norms. We have quite a few examples of symmetric anti-norms on $\mathbb{M}_n^+$. The following are among important examples [6, 7].

Example 2.1. We write $\lambda_j^{\uparrow}(A)$, $j = 1, \dots, n$, for the eigenvalues of $A \in \mathbb{M}_n^+$ in increasing order with counting multiplicities, and similarly $\lambda_j^{\downarrow}(A)$, $j = 1, \dots, n$, for the eigenvalues of A in decreasing order.

(i) For $k = 1, \dots, n$ the *Ky Fan k-anti-norm* is

$$\|A\|_{\{k\}} := \sum_{j=1}^{k} \lambda_j^{\uparrow}(A).$$

This is the anti-norm version of the *Ky Fan k-norm* $\|A\|_{(k)} := \sum_{j=1}^{k} \lambda_j^{\downarrow}(A)$. It is remarkable that the trace functional $\mathrm{Tr}\, A = \|A\|_{(n)} = \|A\|_{\{n\}}$ is a symmetric norm and a symmetric anti-norm simultaneously.

(ii) The *Schatten quasi-norm* $\|A\|_p := \{\mathrm{Tr}\,(A^p)\}^{1/p}$ when $0 < p < 1$ is a symmetric anti-norm. For $p > 0$ so is the *negative Schatten anti-norm* $\|A\|_{-p} := \{\mathrm{Tr}\,(A^{-p})\}^{-1/p}$ (defined to be 0 unless A is invertible).

(iii) For $k = 1, \dots, n$ the functional of Minkowski type

$$\Delta_k(A) := \left\{ \prod_{j=1}^{k} \lambda_j^{\uparrow}(A) \right\}^{1/k}.$$

is a symmetric anti-norm. In particular, $\Delta_n(A) = \det^{1/n} A$ is the *Minkowski functional.*

In this section we deal with convexity or concavity properties for symmetric norm or anti-norm functions of the form $\|\{\Phi(A^p)\,\sigma\,\Psi(B^q)\}^s\|$ or $\|\{\Phi(A^p)\,\sigma\,\Psi(B^q)\}^s\|_!$ involving an operator mean σ in the Kubo-Ando sense [16]. As in the previous section we assume that $(p, q) \neq (0, 0)$ and $s \neq 0$.

Theorem 2.2. *Let $n, m, l \in \mathbb{N}$. Let $\Phi : \mathbb{M}_n \to \mathbb{M}_l$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_l$ be strictly positive linear maps, and σ be any operator mean in the Kubo-Ando sense. Assume that either $0 \le p, q \le 1$ and $0 < s \le 1/\max\{p, q\}$, or $-1 \le p, q \le 0$ and $1/\min\{p, q\} \le s < 0$.*

(1) *For every symmetric anti-norm $\|\cdot\|_!$ on $\mathbb{M}_l^+$ the function*

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \|\{\Phi(A^p)\,\sigma\,\Psi(B^q)\}^s\|_!$$

is jointly concave.

(2) *For every symmetric norm $\|\cdot\|$ on $\mathbb{M}_l$ the function*

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \|\{\Phi(A^p)\,\sigma\,\Psi(B^q)\}^{-s}\|^{-1}$$

is jointly concave, and hence the function

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \|\{\Phi(A^p)\,\sigma\,\Psi(B^q)\}^{-s}\|$$

is jointly convex.

To prove joint concavity of the anti-norm function, we shall use the particular case with the trace function, which we first show as a lemma. Even this trace function case extends [13, Theorem 4.3].

Lemma 2.3. *Let Φ, Ψ and σ be as in Theorem 2.2. Under the same assumptions of p, q and s as in Theorem 2.2, the function*

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \mathrm{Tr}\,\{\Phi(A^p)\,\sigma\,\Psi(B^q)\}^s$$

is jointly concave.

Proof. First, assume that $0 < p, q \le 1$ and $0 < s \le 1$. Let $A_1, A_2 \in \mathbb{P}_n$ and $B_1, B_2 \in \mathbb{P}_m$. By monotonicity and joint concavity of σ we have

$$\begin{aligned}
&\Phi\biggl(\Bigl(\frac{A_1 + A_2}{2}\Bigr)^p\biggr)\,\sigma\,\Psi\biggl(\Bigl(\frac{B_1 + B_2}{2}\Bigr)^q\biggr) \\
&\qquad \ge \biggl(\frac{\Phi(A_1^p) + \Phi(A_2^p)}{2}\biggr)\,\sigma\,\biggl(\frac{\Psi(B_1^q) + \Psi(B_2^q)}{2}\biggr) \\
&\qquad \ge \frac{1}{2}\bigl\{(\Phi(A_1^p)\,\sigma\,\Psi(B_1^q)) + (\Phi(A_2^p)\,\sigma\,\Psi(B_2^q))\bigr\}
\end{aligned} \tag{2.1}$$

and so

$$\begin{aligned}
&\biggl\{\Phi\biggl(\Bigl(\frac{A_1 + A_2}{2}\Bigr)^p\biggr)\,\sigma\,\Psi\biggl(\Bigl(\frac{B_1 + B_2}{2}\Bigr)^q\biggr)\biggr\}^s \\
&\qquad \ge \frac{1}{2}\bigl\{(\Phi(A_1^p)\,\sigma\,\Psi(B_1^q))^s + (\Phi(A_2^p)\,\sigma\,\Psi(B_2^q))^s\bigr\}.
\end{aligned}$$

Secondly, assume that $0 < p, q \le 1$ and $1 \le s \le 1/\max\{p, q\}$. By taking account of the transposed mean $A\,\sigma'\,B := B\,\sigma\,A$ [16], we may further assume that $q \le p$ and so $1 \le s \le 1/p$. We show that the assertion in this case follows from that in the more particular case $q = p$ and $s = 1/p$. Indeed, let $p' := 1/s \in [p, 1]$. Since

$$\begin{aligned}
\Phi\biggl(\Bigl(\frac{A_1 + A_2}{2}\Bigr)^p\biggr) &\ge \Phi\biggl(\Bigl(\frac{A_1^{p/p'} + A_2^{p/p'}}{2}\Bigr)^{p'}\biggr), \\
\Psi\biggl(\Bigl(\frac{B_1 + B_2}{2}\Bigr)^q\biggr) &\ge \Psi\biggl(\Bigl(\frac{B_1^{q/p'} + B_2^{q/p'}}{2}\Bigr)^{p'}\biggr),
\end{aligned}$$

the joint concavity assertion for p, q both replaced with p' implies that

$$\begin{aligned}
&\mathrm{Tr}\biggl\{\Phi\biggl(\Bigl(\frac{A_1 + A_2}{2}\Bigr)^p\biggr)\,\sigma\,\Psi\biggl(\Bigl(\frac{B_1 + B_2}{2}\Bigr)^q\biggr)\biggr\}^s \\
&\quad \ge \mathrm{Tr}\biggl\{\Phi\biggl(\Bigl(\frac{A_1^{p/p'} + A_2^{p/p'}}{2}\Bigr)^{p'}\biggr)\,\sigma\,\Psi\biggl(\Bigl(\frac{B_1^{q/p'} + B_2^{q/p'}}{2}\Bigr)^{p'}\biggr)\biggr\}^s \\
&\quad \ge \frac{1}{2}\Bigl(\mathrm{Tr}\,\bigl\{\Phi((A_1^{p/p'})^{p'})\,\sigma\,\Psi((B_1^{q/p'})^{p'})\bigr\}^s + \mathrm{Tr}\,\bigl\{\Phi((A_2^{p/p'})^{p'})\,\sigma\,\Psi((B_2^{q/p'})^{p'})\bigr\}^s\Bigr)
\end{aligned}$$

$$= \frac{1}{2}\Big(\mathrm{Tr}\,\big\{\Phi(A_1^p)\,\sigma\,\Psi(B_1^q)\big\}^s + \mathrm{Tr}\,\big\{\Phi(A_2^p)\,\sigma\,\Psi(B_2^q)\big\}^s\Big).$$

Hence the proof is reduced to joint concavity of $\mathrm{Tr}\,\{\Phi(A^p)\,\sigma\,\Psi(B^p)\}^{1/p}$ when $0 < p \le 1$. Now the proof can be done by a slight modification of that of [13, Theorem 4.3] (also that of Theorem 1.1 (1) above). We omit the details.

Finally, to treat the case where $-1 \le p, q \le 0$ and $1/\min\{p, q\} \le s < 0$, we can use the same technique as in the last paragraph of the proof of Theorem 1.1 (1). Consider $\hat\Phi, \hat\Psi$ as given there and the adjoint operator mean $A\,\sigma^*\,B := (A^{-1}\,\sigma\,B^{-1})^{-1}$ [16]; then we notice that

$$\{\hat\Phi(A^p)\,\sigma^*\,\hat\Psi(B^q)\}^s = \{\Phi(A^{-p})\,\sigma\,\Psi(B^{-q})\}^{-s}.$$

Hence it remains to see that the joint concavity assertion for Φ, Ψ is valid for $\hat\Phi, \hat\Psi$ too. Indeed, inequality (2.1) follows from Lemma 2.4 below and all other arguments in the above proof can be repeated only using monotonicity of $\hat\Phi, \hat\Psi$. Moreover, the proof of [13, Theorem 4.3] can easily be modified to obtain the joint concavity of $\mathrm{Tr}\,\{\hat\Phi(A^p)\,\sigma\,\hat\Psi(B^p)\}^{1/p}$ when $0 < p \le 1$. $\square$

Lemma 2.4. *Let $\Phi : \mathbb{M}_n \to \mathbb{M}_l$ be a strictly positive linear map. If $0 \le p \le 1$, then the function $A \in \mathbb{P}_n \mapsto \Phi(A^{-p})^{-1} = \hat\Phi(A^p)$ is operator concave.*

Proof. Recall [1, Corollary 3.2] that $A \in \mathbb{P}_n \mapsto \Phi(A^{-1})^{-1}$ is operator concave. For every $A, B \in \mathbb{P}_n$, since $((A + B)/2)^{-p} \le ((A^p + B^p)/2)^{-1}$, we have

$$\Phi\bigg(\Big(\frac{A+B}{2}\Big)^{-p}\bigg)^{-1} \ge \Phi\bigg(\Big(\frac{A^p+B^p}{2}\Big)^{-1}\bigg)^{-1} \ge \frac{\Phi(A^{-p})^{-1} + \Phi(B^{-p})^{-1}}{2}.$$

$\square$

Proof of Theorem 2.2. (1) For every $A_1, A_2 \in \mathbb{M}_n^+$ and $B_1, B_2 \in \mathbb{M}_m^+$ and every Ky Fan k-anti-norm $\|\cdot\|_{\{k\}}$, $1 \le k \le l$, there exists a rank k projection E commuting with $\Phi(((A_1 + A_2)/2)^p)\,\sigma\,\Psi(((B_1 + B_2)/2)^q)$ such that

$$\begin{aligned}
&\left\|\left\{\Phi\bigg(\Big(\frac{A_1+A_2}{2}\Big)^p\bigg)\,\sigma\,\Psi\bigg(\Big(\frac{B_1+B_2}{2}\Big)^q\bigg)\right\}^s\right\|_{\{k\}} \\
&= \mathrm{Tr}\left\{E\bigg(\Phi\bigg(\Big(\frac{A_1+A_2}{2}\Big)^p\bigg)\,\sigma\,\Psi\bigg(\Big(\frac{B_1+B_2}{2}\Big)^q\bigg)\bigg)E\right\}^s \\
&= \lim_{\varepsilon\searrow 0}\mathrm{Tr}\left\{(E+\varepsilon I_l)\bigg(\Phi\bigg(\Big(\frac{A_1+A_2}{2}\Big)^p\bigg)\,\sigma\,\Psi\bigg(\Big(\frac{B_1+B_2}{2}\Big)^q\bigg)\bigg)(E+\varepsilon I_l)\right\}^s \\
&= \lim_{\varepsilon\searrow 0}\mathrm{Tr}\left\{\bigg((E+\varepsilon I_l)\bigg(\Phi\bigg(\Big(\frac{A_1+A_2}{2}\Big)^p\bigg)(E+\varepsilon I_l)\bigg)\right. \\
&\qquad\qquad \left.\sigma\,\bigg((E+\varepsilon I_l)\Psi\bigg(\Big(\frac{B_1+B_2}{2}\Big)^q\bigg)\bigg)(E+\varepsilon I_l)\bigg)\right\}^s
\end{aligned} \tag{2.2}$$

due to the transformer equality for σ [16]. Apply Lemma 2.3 to the positive linear maps $(E+\varepsilon I_l)\Phi(\cdot)(E+\varepsilon I_l)$ and $(E+\varepsilon I_l)\Psi(\cdot)(E+\varepsilon I_l)$ to obtain

$$
\begin{aligned}
\operatorname{Tr}\Bigg\{&\Bigg((E+\varepsilon I_l)\Bigg(\Phi\Bigg(\Bigg(\frac{A_1+A_2}{2}\Bigg)^p\Bigg)(E+\varepsilon I_l)\Bigg) \\
&\qquad\qquad \sigma\Bigg((E+\varepsilon I_l)\Psi\Bigg(\Bigg(\frac{B_1+B_2}{2}\Bigg)^q\Bigg)\Bigg)(E+\varepsilon I_l)\Bigg)\Bigg\}^s \\
&\ge \frac{1}{2}\Big(\operatorname{Tr}\big\{((E+\varepsilon I_l)\Phi(A_1^p)(E+\varepsilon I_l))\,\sigma\,((E+\varepsilon I_l)\Psi(B_1^q)(E+\varepsilon I_l))\big\}^s \\
&\qquad\qquad + \operatorname{Tr}\big\{((E+\varepsilon I_l)\Phi(A_2^p)(E+\varepsilon I_l))\,\sigma\,((E+\varepsilon I_l)\Psi(B_2^q)(E+\varepsilon I_l))\big\}^s\Big) \\
&= \frac{1}{2}\Big(\operatorname{Tr}\big\{(E+\varepsilon I_l)(\Phi(A_1^p)\,\sigma\,\Psi(B_1^q))(E+\varepsilon I_l)\big\}^s \\
&\qquad\qquad + \operatorname{Tr}\big\{(E+\varepsilon I_l)(\Phi(A_2^p)\,\sigma\,\Psi(B_2^q))(E+\varepsilon I_l)\big\}^s\Big) \\
&\longrightarrow \frac{1}{2}\Big(\operatorname{Tr}\big\{E(\Phi(A_1^p)\,\sigma\,\Psi(B_1^q))E\big\}^s + \operatorname{Tr}\big\{E(\Phi(A_2^p)\,\sigma\,\Psi(B_2^q))E\big\}^s\Big) \quad \text{as } \varepsilon \searrow 0.
\end{aligned} \tag{2.3}
$$

For each $C \in \mathbb{M}_l^+$, besides $\lambda_j^\uparrow(C)$ and $\lambda_j^\downarrow(C)$, $1 \le j \le l$, we write $\lambda_j^\uparrow(ECE)$ and $\lambda_j^\downarrow(ECE)$, $1 \le j \le k$, for the eigenvalues (in increasing and decreasing order, respectively) of $ECE|_{E\mathbb{C}^l}$ regarded as an element of $\mathbb{M}_k^+$. Note that

$$
\lambda_j^\uparrow(ECE) \ge \lambda_j^\uparrow(C), \quad \lambda_j^\downarrow(ECE) \le \lambda_j^\downarrow(C), \qquad j = 1,\dots,k. \tag{2.4}
$$

Hence we have for $s > 0$,

$$
\begin{aligned}
\operatorname{Tr}\big\{E(\Phi(A_1^p)\,\sigma\,\Psi(B_1^q))E\big\}^s &= \sum_{j=1}^k \big\{\lambda_j^\uparrow(E(\Phi(A_1^p)\,\sigma\,\Psi(B_1^q))E)\big\}^s \\
&\ge \sum_{j=1}^k \big\{\lambda_j^\uparrow(\Phi(A_1^p)\,\sigma\,\Psi(B_1^q))\big\}^s = \|\{\Phi(A_1^p)\,\sigma\,\Psi(B_1^q)\}^s\|_{\{k\}},
\end{aligned}
$$

and the same inequality follows for $s < 0$ as well (by replacing $\lambda_j^\uparrow$ with $\lambda_j^\downarrow$). Similarly

$$
\operatorname{Tr}\big\{E(\Phi(A_2^p)\,\sigma\,\Psi(B_2^q))E\big\}^s \ge \|\{\Phi(A_2^p)\,\sigma\,\Psi(B_2^q)\}^s\|_{\{k\}}.
$$

Combining these with (2.2) and (2.3) yields that

$$
\begin{aligned}
&\Bigg\|\Bigg\{\Phi\Bigg(\Bigg(\frac{A_1+A_2}{2}\Bigg)^p\Bigg)\,\sigma\,\Psi\Bigg(\Bigg(\frac{B_1+B_2}{2}\Bigg)^q\Bigg)\Bigg\}^s\Bigg\|_{\{k\}} \\
&\quad \ge \frac{1}{2}\Big(\big\|\{\Phi(A_1^p)\,\sigma\,\Psi(B_1^q)\}^s\big\|_{\{k\}} + \big\|\{\Phi(A_2^p)\,\sigma\,\Psi(B_2^q)\}^s\big\|_{\{k\}}\Big) \\
&\quad = \frac{1}{2}\big\|\big(\{\Phi(A_1^p)\,\sigma\,\Psi(B_1^q)\}^s\big)^\uparrow + \big(\{\Phi(A_1^p)\,\sigma\,\Psi(B_1^q)\}^s\big)^\uparrow\big\|_{\{k\}},
\end{aligned}
$$

where $C^{\uparrow}$ for $C \in \mathbb{M}_l^+$ means the diagonal matrix $\mathrm{diag}(\lambda_1^{\uparrow}(C), \dots, \lambda_l^{\uparrow}(C))$. Therefore, for any anti-norm,

$$\begin{aligned}
&\left\| \left\{ \Phi\left(\left(\frac{A_1 + A_2}{2} \right)^p \right) \sigma\, \Psi\left(\left(\frac{B_1 + B_2}{2} \right)^q \right) \right\}^s \right\|_! \\
&\qquad \ge \frac{1}{2} \left\| \left(\{\Phi(A_1^p)\, \sigma\, \Psi(B_1^q)\}^s \right)^{\uparrow} + \left(\{\Phi(A_1^p)\, \sigma\, \Psi(B_1^q)\}^s \right)^{\uparrow} \right\|_! \\
&\qquad \ge \frac{1}{2} \left(\left\| \{\Phi(A_1^p)\, \sigma\, \Psi(B_1^q)\}^s \right\|_! + \left\| \{\Phi(A_1^p)\, \sigma\, \Psi(B_1^q)\}^s \right\|_! \right)
\end{aligned}$$

by the Ky Fan dominance principle [6, Lemma 4.2], the superadditivity property, and unitary conjugation invariance of anti-norms.

(2) is immediately seen by applying (1) to the derived anti-norm $\|A^{-1}\|^{-1}$ for $A \in \mathbb{P}_l$ [7, Proposition 4.6]. ☐

From the particular case of Theorem 2.2 where $\Phi = \Psi = \mathrm{id}$, we have

Corollary 2.5. *Let σ be any operator mean. Assume that either $0 \le p, q \le 1$ and $0 < s \le 1/\max\{p, q\}$, or $-1 \le p, q \le 0$ and $1/\min\{p, q\} \le s < 0$.*

(1) *For any symmetric anti-norm $\|\cdot\|_!$ on $\mathbb{M}_n^+$, $\|(A^p\, \sigma\, B^q)^s\|_!$ is jointly concave in $A, B \in \mathbb{P}_n$.*

(2) *For any symmetric norm $\|\cdot\|$ on $\mathbb{M}_n$, $\|(A^p\, \sigma\, B^q)^{-s}\|^{-1}$ is jointly concave in $A, B \in \mathbb{P}_n$, and hence $\|(A^p\, \sigma\, B^q)^{-s}\|$ is jointly convex in $A, B \in \mathbb{P}_n$.*

In [8] Carlen and Lieb proved that the Minkowski type trace function

$$(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_n^+ \longmapsto \mathrm{Tr}\,(A^p + B^p)^{1/p} \tag{2.5}$$

is jointly concave if $0 < p \le 1$, jointly convex if $p = 2$, and not jointly convex (also not jointly concave) if $p > 2$. The latter assertions when $p = 2$ and when $p > 2$ were also shown in [2], and the former when $0 < p \le 1$ was a bit generalized in [13, Theorem 2.1]. Bekjan [4] later treated joint concavity/convexity of trace functions complementing (2.5) and proved that when $0 < p \le 1$, $\mathrm{Tr}\,(A^{-p} + B^{-p})^{-1/p}$ is jointly concave in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n$, and $\mathrm{Tr}\,(A^{-p} + B^{-p})^{1/p}$ and $\mathrm{Tr}\,(A^p + B^p)^{-1/p}$ are jointly convex in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_n$. Furthermore, in the second paper [9] of the same title, Carlen and Lieb affirmatively settled the conjecture that (2.5) is jointly convex if $1 \le p \le 2$.

The above mentioned trace inequalities in [4] are generalized by the following special case of Theorem 2.2 where σ is the arithmetic mean and both $\|\cdot\|$ and $\|\cdot\|_!$ are the trace functional.

Corollary 2.6. *Let Φ and Ψ be as in Theorem 2.2. Under the same assumption of p, q and s as in Theorem 2.2, $\mathrm{Tr}\,\{\Phi(A^p) + \Psi(B^q)\}^s$ and $\left(\mathrm{Tr}\,\{\Phi(A^p) + \Psi(B^q)\}^{-s}\right)^{-1}$ are jointly concave in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_m$, and hence $\mathrm{Tr}\,\{\Phi(A^p) + \Psi(B^q)\}^{-s}$ is jointly convex in $(A, B) \in \mathbb{P}_n \times \mathbb{P}_m$.*

Corollary 2.7. *Assume that $\Phi : \mathbb{M}_n \to \mathbb{M}_l$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_l$ are positive linear maps such that $\Phi(I_n) + \Psi(I_m) = I_l$.*

(1) *For any symmetric anti-norm $\|\cdot\|_!$ on $\mathbb{M}_l^+$, the function*

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \|\exp\{\Phi(\log A) + \Psi(\log B)\}\|_!$$

is jointly concave.

(2) *For any symmetric norm $\|\cdot\|$ on $\mathbb{M}_l$, the function*

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \|\exp\{\Phi(-\log A) + \Psi(-\log B)\}\|^{-1}$$

is jointly concave, and hence the function

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \|\exp\{\Phi(-\log A) + \Psi(-\log B)\}\|$$

is jointly convex.

Indeed, we may assume by continuity that Φ and Ψ are strictly positive linear maps. Then the assertion (1) follows from Theorem 2.2 (1) with σ the arithmetic mean since

$$\lim_{p \searrow 0} \{\Phi(A^p) + \Psi(B^p)\}^{1/p} = \exp\{\Phi(\log A) + \Psi(\log B)\}.$$

The assertion (2) is a consequence of (1) and [7, Proposition 4.6] as before. By choosing the Minkowski functional $\det^{1/n}$ as $\|\cdot\|_!$, the above (1) implies that

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \mapsto \exp\{\tau(\Phi(\log A) + \Psi(\log B))\}$$

is jointly concave, where τ is the normalized trace on $\mathbb{M}_l$. This is similar to [6, Theorem 5.2].

3 Norm functions of Epstein type

In this section we deal with convexity or concavity properties for symmetric norm or anti-norm functions of the form $\|\Phi(A^p)^s\|$ or $\|\Phi(A^p)^s\|_!$ where Φ is a (strictly) positive linear map. In particular, when $\Phi(A) := X^*AX$ and $\|\cdot\|$ (or $\|\cdot\|_!$) is the trace functional, the function is

$$A \in \mathbb{P}_n \longmapsto \mathrm{Tr}\,(X^*A^pX)^s, \tag{3.1}$$

whose concavity when $0 < p \le 1$ and $s = 1/p$ was established by Epstein [10] (see [13] for some generalizations). It was proved in [9] that the function (3.1) is, for any $X \in \mathbb{M}_n$, convex if $1 \le p \le 2$ and $s \ge 1/p$, concave if $0 < p \le 1$ and $1 \le s \le 1/p$, and

neither convex nor concave if $p > 2$. It was also pointed out there that joint concavity of (2.5) when $0 < p \le 1$ is easily seen from Epstein's concavity [10] (see the first paragraph of the next section) since

$$\operatorname{Tr}\left(\begin{bmatrix} I & 0 \\ I & 0 \end{bmatrix}^* \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^p \begin{bmatrix} I & 0 \\ I & 0 \end{bmatrix}\right)^{1/p} = \operatorname{Tr}(A^p + B^p)^{1/p}. \tag{3.2}$$

In the rest of the section we assume that p and s are non-zero; otherwise, the assertion is trivial.

Theorem 3.1. *Let $n, m \in \mathbb{N}$, and $\Phi : \mathbb{M}_n \to \mathbb{M}_m$ be a strictly positive linear map.*

(1) *Let $\|\cdot\|_!$ be any symmetric anti-norm on $\mathbb{M}_m^+$. If either $0 < p \le 1$ and $0 < s \le 1/p$, or $-1 \le p < 0$ and $1/p \le s < 0$, then $A \in \mathbb{P}_n \mapsto \|\Phi(A^p)^s\|_!$ is concave.*

(2) *Let $\|\cdot\|$ be any symmetric norm on $\mathbb{M}_m$. If either $0 < p \le 1$ and $0 < s \le 1/p$, or $-1 \le p < 0$ and $1/p \le s < 0$, then $A \in \mathbb{P}_n \mapsto \|\Phi(A^p)^{-s}\|^{-1}$ is concave. Moreover, $A \in \mathbb{P}_n \mapsto \|\Phi(A^p)^s\|$ is convex if one of the following three conditions is satisfied:*

$$\begin{cases} -1 \le p < 0 \text{ and } s > 0, \\ 0 < p \le 1 \text{ and } s < 0, \\ 1 \le p \le 2 \text{ and } s \ge 1. \end{cases} \tag{3.3}$$

Proof. (1) and the first assertion of (2) are included in Theorem 2.2 as special case where $B = A$, $\Psi = \Phi$ and $q = p$. For the second assertion of (2) it remains to show that $A \in \mathbb{P}_n \mapsto \|\Phi(A^p)^s\|$ is convex when $p \in [-1, 0) \cup [1, 2]$ and $s \ge 1$ and when $p \in (0, 1]$ and $s < 0$. This can easily be verified as follows: For the latter case, since x^p is operator concave, we have $\Phi(((A + B)/2)^p) \ge (\Phi(A^p) + \Phi(B^p))/2$ for $A, B \in \mathbb{P}_n$. Hence we need to show that $A \in \mathbb{P}_n \mapsto \|A^s\|$ is decreasing and convex when $s < 0$. If $A \le B$, then $(A^s)^\downarrow \ge (B^s)^\downarrow$ so that $\|A^s\| \ge \|B^s\|$. For $A, B \in \mathbb{P}_n$, from the Ky Fan majorization $((A + B)/2)^\downarrow \prec (A^\downarrow + B^\downarrow)/2$ we have

$$\begin{aligned} \left(\left(\frac{A+B}{2}\right)^s\right)^\uparrow &= \left(\left(\frac{A+B}{2}\right)^\downarrow\right)^s \prec_w \left(\frac{A^\downarrow + B^\downarrow}{2}\right)^s \\ &\le \frac{(A^\downarrow)^s + (B^\downarrow)^s}{2} = \frac{(A^s)^\uparrow + (B^s)^\uparrow}{2} \end{aligned}$$

so that

$$\left\|\left(\frac{A+B}{2}\right)^s\right\| \le \left\|\frac{(A^s)^\uparrow + (B^s)^\uparrow}{2}\right\| \le \frac{\|A^s\| + \|B^s\|}{2}.$$

The former case is similarly shown. ∎

The above theorem does not cover the convexity assertion in [9] for (3.1) when $1 \le p \le 2$ and $1/p \le s < 1$. But we can extend this by using the method in [9] itself in the following way. Here, we assume a stronger assumption of Φ being completely positive (CP).

Theorem 3.2. *Let $n, m \in \mathbb{N}$, and $\Phi : \mathbb{M}_n \to \mathbb{M}_m$ be a CP linear map. Let $\|\cdot\|$ be any symmetric norm on $\mathbb{M}_m$. If $1 \le p \le 2$ and $s \ge 1/p$, then $A \in \mathbb{M}_n^+ \mapsto \|\Phi(A^p)^s\|$ is convex.*

Proof. By Theorem 3.1 (2) we may assume that $1/p \le s \le 1$. First, we prove the trace function case. This part of the proof is a slight modification of that in [9]. Let $r := 1/s$ so that $1 \le r \le p \le 2$. By [9, Lemma 2.2] we have

$$\mathrm{Tr}\,\Phi(A^p)^{1/r} = \frac{1}{r} \inf_{B \in \mathbb{P}_m} \mathrm{Tr}\,\{\Phi(A^p)B^{1-r} + (r-1)B\}.$$

Hence it suffices to show that

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \mathrm{Tr}\,\Phi(A^p)B^{1-r}$$

is jointly convex. Since $-1 \le 1 - r \le 0$ and $1 - (1 - r) \le p \le 2$, it follows from [1, Corollary 6.3] that $(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \mapsto A^p \otimes B^{1-r} \in \mathbb{M}_m \otimes \mathbb{M}_m$ is jointly convex. Hence so is

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto (\Phi \otimes \mathrm{id})(A^p \otimes B^{1-r}) = \Phi(A^p) \otimes B^{1-r} \in \mathbb{M}_m \times \mathbb{M}_m,$$

since $\Phi \otimes \mathrm{id}$ is positive by the CP assumption of Φ. This implies the assertion for the trace function.

Next, we extend the result to the symmetric norm function. Since the proof is similar to (and easier than) that of Theorem 2.2 (1), we only sketch it. For every $A, B \in \mathbb{M}_n^+$ and every Ky Fan k-norm $\|\cdot\|_{(k)}$, $1 \le k \le m$, there exists a rank k projection E commuting with $\Phi(((A + B)/2)^p)$ such that

$$\left\| \Phi\left(\left(\frac{A+B}{2} \right)^p \right)^s \right\|_{(k)} = \mathrm{Tr}\left\{ E\Phi\left(\left(\frac{A+B}{2} \right)^p \right) E \right\}^s.$$

Applying the trace function case to the CP linear map $E\Phi(\cdot)E$ we have

$$\mathrm{Tr}\left\{ E\Phi\left(\left(\frac{A+B}{2} \right)^p \right) E \right\}^s \le \frac{1}{2}\big(\mathrm{Tr}\,\{E\Phi(A^p)E\}^s + \mathrm{Tr}\,\{E\Phi(B^p)E\}^s\big).$$

Using the second inequality of (2.4) we have $\mathrm{Tr}\left\{E\Phi(A^p)E\right\}^s \le \|\Phi(A^p)^s\|_{(k)}$ and $\mathrm{Tr}\,\{E(\Phi(B^p)E\}^s \le \|\Phi(B^p)^s\|_{(k)}$ so that

$$\left\| \Phi\left(\left(\frac{A+B}{2} \right)^p \right)^s \right\|_{(k)} \le \frac{1}{2}\big(\|\Phi(A^p)^s\|_{(k)} + \|\Phi(B^p)^s\|_{(k)}\big),$$

which implies the desired convexity assertion. ∎

Complementing to Corollary 2.6 we give

Corollary 3.3. *Let $n, m, l \in \mathbb{N}$. Let $\Phi : \mathbb{M}_n \to \mathbb{M}_l$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_l$ be CP linear maps. Assume that $1 \le p \le 2$ and $s \ge 1/p$. Then for any symmetric norm $\|\cdot\|$ on $\mathbb{M}_l$, the function*

$$(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_m^+ \longmapsto \|\{\Phi(A^p) + \Psi(B^p)\}^s\|$$

is jointly convex, and in particular so is $\mathrm{Tr}\,\{\Phi(A^p) + \Psi(B^p)\}^s$ *in* $(A, B) \in \mathbb{M}_n^+ \times \mathbb{M}_m^+$.

Proof. The proof is a slight modification of expression (3.2). Consider a linear map $\Theta : \mathbb{M}_{n+m} \to \mathbb{M}_l$ defined by

$$\Theta\left(\begin{bmatrix} A & X \\ Y & B \end{bmatrix}\right) := \Phi(A) + \Psi(B)$$

in the form of block matrices with $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_m$. It is easy to see that Θ is CP. Since

$$\{\Phi(A^p) + \Psi(B^p)\}^s = \Theta\left(\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^p\right)^s, \qquad A \in \mathbb{M}_n^+,\ B \in \mathbb{M}_m^+,$$

the assertion is an immediate consequence of Theorem 3.2. ☐

4 Necessary conditions

In the previous sections we obtained sufficient conditions on the parameters p, q, s (or p, s) for which the relevant matrix trace or norm function is (jointly) concave or convex. The aim of this section is to specify necessary conditions on the parameters for those concavity/convexity properties to hold.

Concerning the necessity direction for (joint) concavity of (1.1) and (3.1) we have

Proposition 4.1. (1) *Assume that $p, s \ne 0$. If $A \in \mathbb{P}_2 \mapsto \mathrm{Tr}\,(X^* A^p X)^s$ is concave for any invertible $X \in \mathbb{M}_2$, then either $0 < p \le 1$ and $0 < s \le 1/p$, or $-1 \le p < 0$ and $1/p \le s < 0$.*

(2) *Assume that $(p, q) \ne (0, 0)$ and $s \ne 0$. If $(A, B) \in \mathbb{P}_2 \times \mathbb{P}_2 \mapsto \mathrm{Tr}\,(A^{p/2} B^q A^{p/2})^s$ is jointly concave, then either $0 \le p, q \le 1$ and $0 < s \le 1/(p+q)$, or $-1 \le p, q \le 0$ and $1/(p+q) \le s < 0$.*

Proof. (1) First assume that $s > 0$. By assumption, x^{ps} is concave in $x > 0$ so that $0 < ps \le 1$. For every $a, b, \varepsilon > 0$ let $A := \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ and $X_\varepsilon := \begin{bmatrix} 1 & 0 \\ 1 & \varepsilon \end{bmatrix}$; then $\mathrm{Tr}\,(X_\varepsilon^* A^p X_\varepsilon)^s \to (a^p + b^p)^s$ as $\varepsilon \searrow 0$. So $(a^p + b^p)^s$ is concave in $a, b > 0$. Since

$$\frac{d^2}{dx^2}\,(x^p + b)^s = psx^{p-2}(x^p + b)^{s-2}\{(ps - 1)x^p + (p - 1)b\}, \tag{4.1}$$

we must have $(ps-1)x^p + (p-1)b \le 0$ for all $x, b > 0$, which gives $p \le 1$ as well as $ps \le 1$. When $s < 0$, the result follows from the above case since $\mathrm{Tr}\,(X^*A^pX)^s = \mathrm{Tr}\,(X^{-1}A^{-p}(X^{-1})^*)^{-s}$.

(2) As in the proof of (1) it suffices to assume that $s > 0$. By assumption, $x^{(p+q)s}$ is concave in $x > 0$ so that $(p+q)s \le 1$. The assumption also implies that $A \in \mathbb{P}_2 \mapsto \mathrm{Tr}\,(X^*A^pX)^s$ is concave for every invertible $X \in \mathbb{M}_2$, as readily seen by taking the polar decomposition of X. Hence (1) implies that $0 \le p \le 1$. Similarly, $0 \le q \le 1$. $\square$

There is no gap between a necessary condition in Proposition 4.1 (1) and a sufficient condition in Theorem 3.1 (1). This says that Theorem 3.1 (1) is a best possible result. The difference between a necessary condition in Proposition 4.1 (2) and a sufficient condition in Theorem 1.1 (1) is rather small: $0 < s < 1/2$ for $0 \le p, q \le 1$, or $-1/2 < s < 0$ for $-1 \le p, q \le 0$. When restricted to $s = 1$, Proposition 4.1 (2) says that a necessary condition for joint concavity of $(A, B) \in \mathbb{P}_2 \times \mathbb{P}_2 \mapsto \mathrm{Tr}\,A^pB^q$ is that $0 \le p, q \le 1$ and $p + q \le 1$, which is also sufficient for joint convexity of (1.2) for any $X \in \mathbb{M}_n$ (as shown in [17, 1]) and even for (1.1) with $s = 1$.

Remark 4.2. For the case in the gap between conditions of Proposition 4.1 (2) and of Theorem 1.1 (1), the following is worth noting: Assume that $0 < p, q \le 1$ and $0 < s \le 1$. For every positive linear maps $\Phi : \mathbb{M}_n \to \mathbb{M}_l$, $\Psi : \mathbb{M}_m \to \mathbb{M}_l$ and for every $A_1, A_2 \in \mathbb{M}_n^+$, $B_1, B_2 \in \mathbb{M}_m^+$ one has

$$\left\{\Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2}\Psi\left(\left(\frac{B_1+B_2}{2}\right)^q\right)\Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2}\right\}^s$$
$$\ge U\left\{\left(\frac{\Phi(A_1^q)+\Phi(A_2^q)}{2}\right)^{1/2}\left(\frac{\Psi(B_1^p)+\Psi(B_2^p)}{2}\right)\left(\frac{\Phi(A_1^q)+\Phi(A_2^q)}{2}\right)^{1/2}\right\}^s U^*$$

for some unitary $U \in \mathbb{M}_l$. Hence, to settle the case $0 < s < 1/2$ (and $0 \le p, q \le 1$) of Theorem 1.1 (1), we need to prove that $(A, B) \in \mathbb{P}_n^+ \times \mathbb{P}_n^+ \mapsto \mathrm{Tr}\,(A^{1/2}BA^{1/2})^s$ is jointly concave if $0 < s < 1/2$.

Concerning the necessity direction for (joint) convexity of (1.1) and (3.1) we have

Proposition 4.3. (1) *Assume that $p, s \ne 0$. If $A \in \mathbb{P}_4 \mapsto \mathrm{Tr}\,(X^*A^pX)^s$ is convex for every $X \in \mathbb{P}_4$, then one of the following four conditions is satisfied:*

$$\begin{cases} -1 \le p < 0 \text{ and } s > 0, \\ 1 \le p \le 2 \text{ and } s \ge 1/p, \end{cases}$$

and their counterparts where (p, s) is replaced with $(-p, -s)$.

(2) *Assume that* $(p,q)\ne(0,0)$ *and* $s\ne 0$. *If* $(A,B)\in\mathbb{P}_4\times\mathbb{P}_4\mapsto \mathrm{Tr}\,(A^{p/2}B^qA^{p/2})^s$ *is jointly convex, then one of the following six conditions is satisfied:*

$$\begin{cases} -1\le p,q\le 0 \ \textit{and}\ s>0, \\ -1\le p\le 0,\ 1\le q\le 2,\ p+q>0 \ \textit{and}\ s\ge 1/(p+q), \\ 1\le p\le 2,\ -1\le q\le 0,\ p+q>0 \ \textit{and}\ s\ge 1/(p+q), \end{cases}$$

and their counterparts where (p,q,s) *is replaced with* $(-p,-q,-s)$.

Proof. As in the proof of Proposition 4.1 it suffices to assume that $s>0$.

(1) Let $X_\varepsilon := \begin{bmatrix} I_2 & 0 \\ I_2 & \varepsilon I_2 \end{bmatrix}\in\mathbb{M}_4$ for $\varepsilon>0$. For any $A,B\in\mathbb{P}_2$, since

$$\mathrm{Tr}\left(X_\varepsilon^*\begin{bmatrix} A & 0 \\ 0 & B\end{bmatrix}^p X_\varepsilon\right)^s \longrightarrow \mathrm{Tr}\,(A^p+B^p)^s \quad \text{as } \varepsilon\searrow 0,$$

the assumption implies that $(A,B)\in\mathbb{P}_2\times\mathbb{P}_2\mapsto\mathrm{Tr}\,(A^p+B^p)^s$ is jointly convex, so $\varphi_t(A):=\mathrm{Tr}\,(tA^p+B)^s$ is convex in $A\in\mathbb{P}_2$ for any $t>0$ and $B\in\mathbb{P}_2$. The argument below is the same as that in [4] (also [9]) while it is given for completeness. Since

$$\frac{d}{dt}\,\varphi_t(A)\Big|_{t=0} = s\mathrm{Tr}\,B^{s-1}A^p,$$

we notice that

$$\varphi_t(A)=\mathrm{Tr}\,B^s + st\mathrm{Tr}\,B^{s-1}A^p + O(t^2) \quad \text{as } t\searrow 0.$$

Therefore, for $A_1,A_2\in\mathbb{P}_2$ we have

$$\begin{aligned} 0 &\ge \varphi_t\!\left(\frac{A_1+A_2}{2}\right)-\frac{\varphi_r(A_1)+\varphi_r(A_2)}{2} \\ &= st\left\{\mathrm{Tr}\,B^{s-1}\left(\frac{A_1+A_2}{2}\right)^p-\mathrm{Tr}\,B^{s-1}\left(\frac{A_1^p+A_2^p}{2}\right)\right\}+O(t^2) \quad \text{as } t\searrow 0 \end{aligned}$$

so that

$$\mathrm{Tr}\,B^{s-1}\left(\frac{A_1+A_2}{2}\right)^p \le \mathrm{Tr}\,B^{s-1}\left(\frac{A_1^p+A_2^p}{2}\right).$$

When $s\ne 1$, this means that x^p $(x>0)$ is matrix convex of order 2, which is also clear for $s=1$ from the assumption itself. Hence by [12, Proposition 3.1] we must have $-1\le p\le 0$ or $1\le p\le 2$. When $1\le p\le 2$, $ps\ge 1$ since x^{ps} is convex.

(2) Since the assumption here implies that of (1), it follows that either $-1\le p\le 0$, or $1\le p\le 2$ and $s\ge 1/p$. Similarly, either $-1\le q\le 0$, or $1\le q\le 2$ and $s\ge 1/q$. Since $x^{ps}y^{qs}$ is jointly convex in $x,y>0$, computing the Hessian gives

$$pq\{1-(p+q)s\}\ge 0. \tag{4.2}$$

Hence $ps \geq 1$ and $qs \geq 1$ cannot occur simultaneously, so the following three cases are possible (when $s > 0$):

$$\begin{cases} -1 \leq p, q \leq 0, \\ -1 \leq p \leq 0 \text{ and } 1 \leq q \leq 2, \\ 1 \leq p \leq 2 \text{ and } -1 \leq q \leq 0. \end{cases}$$

For the above second case, if $p = 0$ then $s \geq 1/q = 1/(p+q)$, and if $p < 0$ then (4.2) gives $p + q > 0$ and $s \geq 1/(p+q)$. The third case is similar. $\square$

A gap between a necessary condition in Proposition 4.3 (1) and a sufficient condition in (3.3) and Theorem 3.2 together is not so big: only the case $-2 \leq p \leq -1$ and $s \leq 1/p$. But there is quite a big gap between conditions in Proposition 4.3 (2) and in Theorem 1.1 (2). When restricted to $s = 1$, Proposition 4.3 (2) says that a necessary condition for joint convexity of $(A, B) \in \mathbb{P}_4 \times \mathbb{P}_4 \mapsto \operatorname{Tr} A^p B^q$ is that

$$\begin{cases} -1 \leq p, q \leq 0, \\ -1 \leq p \leq 0 \text{ and } 1 - p \leq q \leq 2, \\ -1 \leq q \leq 0 \text{ and } 1 - q \leq p \leq 2, \end{cases}$$

which is exactly a necessary and sufficient condition for joint convexity of (1.2) for any $X \in \mathbb{M}_n$ (see [1, p. 221, Remark (4)]). In this connection see also [9, Lemma 5.2].

5 More discussions

Theorem 1.1 was presented for trace functions while we more generally treated symmetric (anti-) norm functions in Theorems 2.2 and 3.1. So it is desirable to extend Theorem 1.1 to joint concavity/convexity of (anti-) norm functions. The problem can be reduced to joint concavity of the Ky Fan k-anti-norm functions

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \big\| \big\{ \Phi(A^p)^{1/2} \Psi(B^q) \Phi(A^p)^{1/2} \big\}^s \big\|_{\{k\}}$$

and joint convexity of the Ky Fan k-norm functions

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \big\| \big\{ \Phi(A^p)^{1/2} \Psi(B^q) \Phi(A^p)^{1/2} \big\}^s \big\|_{(k)}$$

for $k = 1, \ldots, l$ in the situation of Theorem 1.1. In this section we examine the problem in the special case $k = 1$. As in Section 1 we assume that $(p, q) \neq (0, 0)$ and $s \neq 0$.

Theorem 5.1. *Let $\Phi : \mathbb{M}_n \to \mathbb{M}_l$ and $\Psi : \mathbb{M}_m \to \mathbb{M}_l$ be as in Theorem 1.1.*

(1) *The function*

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \lambda_l \big(\big\{ \Phi(A^p)^{1/2} \Psi(B^q) \Phi(A^p)^{1/2} \big\}^s \big)$$

is jointly concave, where $\lambda_l(C) = \lambda_l^{\downarrow}(C)$ *is the smallest eigenvalue of* $C \in \mathbb{P}_l$*, if one of the following two conditions is satisfied:*

$$\begin{cases} 0 \le p, q \le 1 \ \text{and}\ 0 < s \le 1/(p+q), \\ -1 \le p, q \le 0 \ \text{and}\ 1/(p+q) \le s < 0. \end{cases} \tag{5.1}$$

(2) *The function*

$$(A, B) \in \mathbb{P}_n \times \mathbb{P}_m \longmapsto \big\| \big\{ \Phi(A^p)^{1/2} \Psi(B^q) \Phi(A^p)^{1/2} \big\}^s \big\|_\infty,$$

where $\|\cdot\|_\infty$ *is the operator norm, is jointly convex if one of the following six conditions is satisfied:*

$$\begin{cases} -1 \le p, q \le 0 \ \text{and}\ s > 0, \\ -1 \le p \le 0,\ 1 \le q \le 2,\ p+q > 0 \ \text{and}\ s \ge 1/(p+q), \\ 1 \le p \le 2,\ -1 \le q \le 0,\ p+q > 0 \ \text{and}\ s \ge 1/(p+q), \end{cases} \tag{5.2}$$

and their counterparts where (p, q, s) *is replaced with* $(-p, -q, -s)$*.*

Proof of Theorem 5.1. (1) First, assume that $0 \le p, q \le 1$ and $s = 1/(p+q)$. We show that

$$\begin{aligned} &\lambda_l\Bigg(\Bigg\{ \Phi\Bigg(\bigg(\frac{A_1 + A_2}{2}\bigg)^p \Bigg)^{1/2} \Psi\Bigg(\bigg(\frac{B_1 + B_2}{2}\bigg)^q \Bigg) \Phi\Bigg(\bigg(\frac{A_1 + A_2}{2}\bigg)^{p/2} \Bigg) \Bigg\}^s \Bigg) \\ &\qquad \ge \frac{\lambda_l\big(\{ \Phi(A_1^p)^{1/2} \Psi(B_1^q) \Phi(A_1^p)^{1/2} \}^s \big) + \lambda_l\big(\{ \Phi(A_2^p)^{1/2} \Psi(B_2^q) \Phi(A_2^p)^{1/2} \}^s \big)}{2} \end{aligned} \tag{5.3}$$

for every $A_1, A_2 \in \mathbb{P}_n$ and $B_1, B_2 \in \mathbb{P}_m$. Set

$$\alpha_j := \lambda_l\big(\{ \Phi(A_j^p)^{1/2} \Psi(B_j^q) \Phi(A_j^p)^{1/2} \}^{1/(p+q)} \big), \qquad j = 1, 2.$$

We then have $\Phi(A_j^p)^{1/2} \Psi(B_j^q) \Phi(A_j^p)^{1/2} \ge \alpha_j^{p+q} I$ so that

$$\Psi((\alpha_j^{-1} B_j)^q) \ge \Phi((\alpha_j^{-1} A_j)^p)^{-1}, \qquad j = 1, 2. \tag{5.4}$$

Since x^q is operator concave, we have

$$\bigg(\frac{B_1 + B_2}{\alpha_1 + \alpha_2} \bigg)^q \ge \frac{\alpha_1}{\alpha_1 + \alpha_2} (\alpha_1^{-1} B_1)^q + \frac{\alpha_2}{\alpha_1 + \alpha_2} (\alpha_2^{-1} B_2)^q,$$

which implies that

$$\begin{aligned} \Psi\Bigg(\bigg(\frac{B_1 + B_2}{\alpha_1 + \alpha_2} \bigg)^q \Bigg) &\ge \frac{\alpha_1}{\alpha_1 + \alpha_2} \Phi((\alpha_1^{-1} A_1)^p)^{-1} + \frac{\alpha_2}{\alpha_1 + \alpha_2} \Phi((\alpha_2^{-1} A_2)^p)^{-1} \\ &\ge \bigg\{ \Phi\bigg(\frac{\alpha_1}{\alpha_1 + \alpha_2} (\alpha_1^{-1} A_1)^p + \frac{\alpha_2}{\alpha_1 + \alpha_2} (\alpha_2^{-1} A_2)^p \bigg) \bigg\}^{-1} \end{aligned} \tag{5.5}$$

due to (5.4) and operator convexity of x^{-1}. Moreover, the operator concavity of x^p gives

$$\frac{\alpha_1}{\alpha_1+\alpha_2}(\alpha_1^{-1}A_1)^p + \frac{\alpha_2}{\alpha_1+\alpha_2}(\alpha_2^{-1}A_2)^p \le \left(\frac{A_1+A_2}{\alpha_1+\alpha_2}\right)^p.$$

Inserting this into (5.5) we have

$$\Psi\left(\left(\frac{B_1+B_2}{\alpha_1+\alpha_2}\right)^q\right) \ge \Phi\left(\left(\frac{A_1+A_2}{\alpha_1+\alpha_2}\right)^p\right)^{-1} \tag{5.6}$$

so that

$$\Psi\left(\left(\frac{B_1+B_2}{2}\right)^q\right) \ge \left(\frac{\alpha_1+\alpha_2}{2}\right)^{p+q}\Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{-1}.$$

Therefore,

$$\Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2}\Psi\left(\left(\frac{B_1+B_2}{2}\right)^q\right)\Phi\left(\left(\frac{A_1+A_2}{2}\right)^p\right)^{1/2} \ge \left(\frac{\alpha_1+\alpha_2}{2}\right)^{p+q} I,$$

which implies the assertion when $s = 1/(p+q)$. The assertion when $0 < s < 1/(p+q)$ follows by taking the $s(p+q)$-powers of both sides of (5.3) when $s = 1/(p+q)$ and by applying the concavity of $x^{s(p+q)}$.

For the other case where $-1 \le p, q \le 0$ and $1/(p+q) \le s < 0$, we may check (see (1.10)) that the above proof for the first case can be performed with $\hat{\Phi}$ and $\hat{\Psi}$ in place of Φ and Ψ. A non-trivial part is to prove (5.6) for $\hat{\Phi}$ and $\hat{\Psi}$ from (5.4) for $\hat{\Phi}$ and $\hat{\Psi}$, which can be done by Lemma 2.4 as follows:

$$\begin{aligned}
\hat{\Psi}\left(\left(\frac{B_1+B_2}{\alpha_1+\alpha_2}\right)^q\right) &\ge \frac{\alpha_1}{\alpha_1+\alpha_2}\hat{\Phi}((\alpha_1^{-1}A_1)^p)^{-1} + \frac{\alpha_2}{\alpha_1+\alpha_2}\hat{\Phi}((\alpha_2^{-1}A_2)^p)^{-1} \\
&= \Phi\left(\frac{\alpha_1}{\alpha_1+\alpha_2}(\alpha_1^{-1}A_1)^{-p} + \frac{\alpha_2}{\alpha_1+\alpha_2}(\alpha_2^{-1}A_2)^{-p}\right) \\
&\ge \Phi\left(\left(\frac{A_1+A_2}{\alpha_1+\alpha_2}\right)^{-p}\right) = \hat{\Phi}\left(\left(\frac{A_1+A_2}{\alpha_1+\alpha_2}\right)^p\right)^{-1}.
\end{aligned}$$

(2) Assume that $s > 0$. When $-1 \le p, q \le 0$ and $0 < s \le -1/(p+q)$, since

$$\left\|\left\{\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\right\}^s\right\|_\infty = \lambda_l\left(\left\{\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\right\}^{-s}\right)^{-1},$$

the assertion is immediate from (1) above. When $-1 \le p, q \le 0$ and $s > -1/(p+q)$, since $-s(p+q) > 1$ and

$$\left\|\left\{\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\right\}^s\right\|_\infty = \left\|\left\{\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\right\}^{-1/(p+q)}\right\|_\infty^{-s(p+q)},$$

the assertion follows from the case $s = -1/(p+q)$. Next, assume that (p, q, s) satisfies the second condition in (5.2). As in the above argument it suffices to show the joint convexity when $s = 1/(p+q)$. Set

$$\alpha_j := \left\|\{\Phi(A_j^p)^{1/2}\Psi(B_j^q)\Phi(A_j^p)^{1/2}\}^{1/(p+q)}\right\|_\infty, \qquad j = 1, 2.$$

Then $\Phi(A_j^p)^{1/2}\Psi(B_j^q)\Phi(A_j^p)^{1/2} \le \alpha_j^{p+q}I$, and by the same argument as in the proof of (1) with use of Lemma 2.4 we have

$$\Phi\biggl(\biggl(\frac{A_1+A_2}{2}\biggr)^p\biggr)^{1/2}\Psi\biggl(\biggl(\frac{B_1+B_2}{2}\biggr)^q\biggr)\Phi\biggl(\biggl(\frac{A_1+A_2}{2}\biggr)^p\biggr)^{1/2} \le \biggl(\frac{\alpha_1+\alpha_2}{2}\biggr)^{p+q}I,$$

which implies the desired joint convexity. Since

$$\bigl\|\bigl\{\Phi(A^p)^{1/2}\Psi(B^q)\Phi(A^p)^{1/2}\bigr\}^s\bigr\|_\infty = \bigl\|\bigl\{\Psi(B^q)^{1/2}\Phi(A^p)\Phi(B^q)^{1/2}\bigr\}^s\bigr\|_\infty,$$

the assertion holds also when (p,q,s) satisfies the third condition in (5.2).

Finally, the above proof can be repeated with $\hat\Phi$ and $\hat\Psi$ in place of Φ and Ψ (while we omit the details), which shows the assertion under the other three conditions where $s<0$. ☐

By Theorems 1.1 and 5.1 we see that Theorem 1.1 can be extended to symmetric (anti-) norm functions in particular where $l=2$ (under the same assumption for each of (1) and (2) of Theorem 1.1).

Although the next theorem is concerned with functions of the particular form with $\Phi=\Psi=\mathrm{id}$ in Theorem 5.1 (2) under $p,q\ne 0$ (stronger than $(p,q)\ne(0,0)$), it has an advantage that the condition on the parameters is a necessary and sufficient condition. The theorem indeed says that the conditions on p, q and s in Theorem 5.1 are best possible except the case where $p=0$ or $q=0$. When $q=0$ (and $p\ne 0$), the concavity/convexity properties in Theorem 5.1 are reduced to concavity of $A\in\mathbb{P}_n\mapsto\lambda_m(\Phi(A^p)^s)$ and convexity of $A\in\mathbb{P}_n\mapsto\|\Phi(A^p)^s\|_\infty$ for every strictly positive linear map $\Phi:\mathbb{M}_n\to\mathbb{M}_m$, which are special cases of the concavity/convexity properties of Theorem 3.1.

Theorem 5.2. *Assume that p, q and s are non-zero.*

(1) *The function*

$$(A,B)\in\mathbb{P}_n\times\mathbb{P}_n\longmapsto\lambda_n((A^{p/2}B^qA^{p/2})^s) \tag{5.7}$$

is jointly concave for every $n\in\mathbb{N}$ if and only if p, q and s satisfy one of the conditions in (5.1)*.*

(2) *The function*

$$(A,B)\in\mathbb{P}_n\times\mathbb{P}_n\longmapsto\|(A^{p/2}B^qA^{p/2})^s\|_\infty \tag{5.8}$$

is jointly convex for every $n\in\mathbb{N}$ if and only if p, q and s satisfy one of the conditions in (5.2) *and their counterparts for $(-p,-q,-s)$ in place of (p,q,s).*

Proof. Obviously, the "if" parts of (1) and (2) are included in those of Theorem 5.1. In the rest we prove the "only if" parts under $p,q,s\ne 0$. As in the proofs of Propositions 4.1 and 4.3 we may assume that $s>0$.

(1) Since

$$\left\{\lambda_n((A^{p/2}B^qA^{p/2})^s)\right\}^{-1} = \left\|(A^{-p/2}B^{-q}A^{-p/2})^s\right\|_\infty,$$

the joint concavity of (5.7) implies the joint convexity of (5.8) for $(-p,-q,s)$. By applying the conclusion of (2) (proved below) to $(-p,-q,s)$ (with $s>0$) we have

$$\begin{cases} 0<p,q\le 1, \\ 0<p\le 1,\ -2\le q\le -1,\ p+q<0 \text{ and } s\ge -1/(p+q), \\ -2\le p\le -1,\ 0<q\le 1,\ p+q<0 \text{ and } s\ge -1/(p+q). \end{cases}$$

However, the joint concavity of (5.7) implies that $ps, qs, (p+q)s \in (0,1]$ and hence $p,q>0$. So the latter two cases in the above are impossible to appear, and (1) is shown.

To prove (2), we first show

Lemma 5.3. *Let $n\in\mathbb{N}$ and assume that $p,q\ne 0$ and $s>0$. If the function* (5.8) *is jointly convex, then*

$$\left(\frac{A^{1/q}+B^{1/q}}{2}\right)^q \le \left(\frac{A^{-1/p}+B^{-1/p}}{2}\right)^{-p}$$

for every $A,B\in\mathbb{P}_n$.

Proof. The assumption means that

$$\left\|\left\{\left(\frac{A_1+A_2}{2}\right)^{p/2}\left(\frac{B_1+B_2}{2}\right)^q\left(\frac{A_1+A_2}{2}\right)^{p/2}\right\}^s\right\|_\infty \le \frac{\|(A_1^{p/2}B_1^qA_1^{p/2})^s\|_\infty + \|(A_2^{p/2}B_2^qA_2^{p/2})^s\|_\infty}{2}$$

for every $A_i,B_i\in\mathbb{P}_n$, $i=1,2$. Since $s>0$, the above inequality implies that if $A_i^{p/2}B_i^qA_i^{p/2}\le I$ for $i=1,2$ then

$$\left(\frac{A_1+A_2}{2}\right)^{p/2}\left(\frac{B_1+B_2}{2}\right)^q\left(\frac{A_1+A_2}{2}\right)^{p/2}\le I,$$

that is, if $B_i^q\le A_i^{-p}$ for $i=1,2$ then $((B_1+B_2)/2)^q\le((A_1+A_2)/2)^{-p}$. In particular, letting $A_i=B_i^{-q/p}$ gives

$$\left(\frac{B_1+B_2}{2}\right)^q \le \left(\frac{B_1^{-q/p}+B_2^{-q/p}}{2}\right)^{-p},$$

which is clearly equivalent to the desired inequality. ∎

The "if" part of the next lemma is rather easy as given in [18] (also [11, Chapter 4]) in a more general form. However, it would be beyond the scope of this paper if we supply several counterexamples to prove the "only if" part. So we leave the details of the proof to a separate paper [3].

Lemma 5.4. *Let $p, q \in \mathbb{R}$. Then the matrix inequality*

$$\left(\frac{A^p + B^p}{2}\right)^{1/p} \le \left(\frac{A^q + B^q}{2}\right)^{1/q},$$

where $((A^p + B^p)/2)^{1/p}$ for $p = 0$ means

$$\lim_{p\to 0}\left(\frac{A^p + B^p}{2}\right)^{1/p} = \exp\left(\frac{\log A + \log B}{2}\right),$$

holds for every $n \in \mathbb{N}$ and every $A, B \in \mathbb{P}_n$ if and only if one of the following is satisfied:

$$\begin{cases} 1 \le p \le q, \\ p \le q \le -1, \\ p \le -1,\ q \ge 1, \\ 1/2 \le p \le 1 \le q, \\ p \le -1 \le q \le -1/2. \end{cases}$$

Proof of Theorem 5.2 (continued). Let $s > 0$ and assume the joint convexity of (5.8). As in the proof of Proposition 4.1 (1) (just replace Tr with $\|\cdot\|_\infty$ and concavity with convexity) we see that $(a^p + b^p)^s$ is convex in $a, b > 0$. By (4.1) we have either $ps \ge 1$ and $p \ge 1$, or $ps < 0$ and $p \le 1$. Similarly, we have either $qs \ge 1$ and $q \ge 1$, or $qs < 0$ and $q \le 1$. Therefore, $p, q \in (-\infty, 0) \cup [1, \infty)$. Moreover, by Lemmas 5.3 and 5.4, one of the following must be satisfied:

(a) $1 \le 1/q \le -1/p$,

(b) $1/q \le -1/p \le -1$,

(c) $1/q \le -1$, $-1/p \ge 1$,

(d) $1/2 \le 1/q \le 1 \le -1/p$,

(e) $1/q \le -1 \le -1/p \le -1/2$.

When $p, q < 0$ and so $1/q < 0 < -1/p$, (c) must hold so that $-1 \le p, q < 0$, which is the first case of (5.2). When $p < 0$ and $q \ge 1$ and so $0 < 1/q \le 1$ and $-1/p > 0$, (d) (or (a) with $1/q = 1$) must hold so that $-1 \le p < 0$ and $1 \le q \le 2$. Now assume by contradiction that $p + q = 0$; hence $p = -1$ and $q = 1$, so $x^{-s}y^s$ is jointly convex in $x, y > 0$. But this is impossible since the Hessian is $-s^2x^{-2s-2}y^{2s-2} < 0$. Therefore, $p + q > 0$. Moreover, since $x^{(p+q)s}$ is convex in $x > 0$, $(p + q)s \ge 1$. So we have the second case of (5.2). When $p \ge 1$ and $q < 0$, we similarly have the third case of (5.2). Finally, the case where $p \ge 1$ and $q \ge 1$ cannot be compatible with any of (a)–(e), so this case does not appear. $\square$

Acknowledgments

The author would like to thank Jean-Christophe Bourin who suggested him how to extend concavity results for trace functions to symmetric anti-norm functions, that is the main idea of this paper.

References

[1] T. Ando, Concavity of certain maps on positive definite matrices and applications to Hadamard Products, *Linear Algebra Appl.* **26** (1979), 203–241.

[2] T. Ando and F. Hiai, Hölder type inequalities for matrices, *Math. Ineq. Appl.* **1** (1998), 1–30.

[3] K. M. R. Audenaert and F. Hiai, On matrix inequalities between the power means: counterexamples, in preparation.

[4] T. N. Bekjan, On joint convexity of trace functions, *Linear Algebra Appl.* **390** (2004), 321–327.

[5] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1996.

[6] J.-C. Bourin and F. Hiai, Norm and anti-norm inequalities for positive semi-definite matrices, *Internat. J. Math.* **22** (2011), 1121–1138.

[7] J.-C. Bourin and F. Hiai, Jensen and Minkowski inequalities for operator means and anti-norms, Preprint, 2011.

[8] E. A. Carlen and E. H. Lieb, A Minkowski type trace inequality and strong subadditivity of quantum entropy, *Advances in the Mathematical Sciences, Amer. Math. Soc. Transl. Ser. 2* **189** (1999), 59–68.

[9] E. A. Carlen and E. H. Lieb, A Minkowski type trace inequality and strong subadditivity of quantum entropy II: convexity and concavity, *Lett. Math. Phys.* **83** (2008), 107–126.

[10] H. Epstein, Remarks on two theorems of E. Lieb, *Comm. Math. Phys.* **31** (1973), 317–325.

[11] T. Furuta, J. Mićić Hot, J. Pečarić and Y. Seo, *Mond-Pečarić Method in Operator Inequalities*, Element, Zagreb, 2005.

[12] F. Hansen and J. Tomiyama, Differential analysis of matrix convex functions II, *J. Inequal. Pure Appl. Math.* **10** (2009), Article 32, 5 pp.

[13] F. Hiai, Concavity of certain matrix trace functions, *Taiwanese J. Math.* **5** (2001), 535–554.

[14] F. Hiai, Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization (GSIS selected lectures), *Interdisciplinary Information Sciences* **16** (2010), 139–248.

[15] A. Jenčová and M. B. Ruskai, A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality, *Rev. Math. Phys.* **22** (2010), 1099–1121.

[16] F. Kubo and T. Ando, Means of positive linear operators, *Math. Ann.* **246** (1980), 205–224.

[17] E. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, *Advances in Math.* **11** (1973), 267–288.

[18] J. Mićić, Z. Pavić and J. Pečarić, Jensen's inequality for operators without operator convexity, *Linear Algebra Appl.* **434** (2011) 1228–1237.